Positivity of transition probabilities of infinite-dimensional diffusion processes on ellipsoids

Oxana Manita

February 9, 2016

Abstract: We consider diffusion processes in Hilbert spaces with constant non-degenerate diffusion operators and show that, under broad assumptions on the drift, the transition probabilities of the process are positive on ellipsoids associated with the diffusion operator. This is an infinite-dimensional analogue of positivity of densities of transition probabilities. Our results apply to diffusions corresponding to stochastic partial differential equations.

Keywords: diffusion process in Hilbert space; SPDE; support of distribution; positive density; mild solution; variational solution; Kolmogorov equation.

Author's address: Faculty of Mechanics and Mathematics, Moscow State University, Moscow 119991, Russia; o.manita@lambda.msu.ru.

1 Introduction

Let us consider the stochastic differential equation (SDE)

$$dX_t = dW_t + (AX_t + F(X_t)dt), \quad X_0 = \eta \tag{1}$$

in a Hilbert space H, where W_t is an H-valued Wiener process with covariance operator Q, having eigenvectors $\{e_i\}_{i\in\mathbb{N}}$ and eigenvalues $\{q_i\}_{i\in\mathbb{N}}$, and the corresponding Kolmogorov equation

$$\partial_t \mu_t = \frac{1}{2} q_i \partial^2_{e_i e_i}\mu_t - \partial_{e_i}(b^i(x)\mu_t), \quad \mu_0 = \mathrm{Law}(\eta) \tag{2}$$

for the distributions μ_t of the diffusion process X_t. Here $b^i = \langle A + F, e_i\rangle$, A is a linear (possibly unbounded) operator and F is some function on H. Equations of such a form correspond to stochastic partial differential equations (SPDEs). In typical cases A is an elliptic differential operator.

It is well-known that if the coefficients of the equation are regular enough and the diffusion matrix is non-degenerate, then the transition probabilities of the finite-dimensional diffusion process have strictly positive densities with respect to Lebesgue measure (see [6]). In the non-degenerate case this property is usually derived from the Harnack inequality or from the Girsanov theorem. Another powerful approach is provided by the seminal result of Strook and Varadhan [21]. In the finite-dimensional case they give a full description of the support of the distribution of the diffusion process

$$dX_t = \sigma(X_t) \circ dW_t + F(X_t)dt, \quad X_0 = x,$$

where the SDE is written in the form of Stratonovich. Namely, they showed that the support of the distribution of X_t coincides with the closure in the space of continuous functions of the set of solutions to the appropriate control problem: the Wiener process is replaced by a smooth path – control – and the SDE turns into an ODE in the Hilbert space). More precisely, they showed that supp $\mathrm{Law}(X_t) = \mathcal{S}_t$, where

$$\mathcal{S}_t = \overline{\{y_t : \, u \text{ is piecewise constant and } \dot{y} = \sigma(y_s)u + F(y_s), \quad y_0 = x\}}. \tag{3}$$

We emphasize that this results doesn't require non-degeneracy of the diffusion (and is interesting mostly in the degenerate case).

However, in the infinite-dimensional case the situation is different. First of all, in the infinite-dimensional case there is no Lebesgue measure. Therefore, we consider the following property: the measure of every open set is strictly positive. In the finite-dimensional case this holds in case of existence of a strictly positive density with respect to Lebesgue measure. Even for the best studied class of measures in the infinite-dimensional spaces – Gaussian measures – this property is not quite trivial (see [7, Theorem 3.5.1]). Positivity on open sets is sometimes called irreducibility of the semigroup corresponding to the diffusion process (irreducibility of the generator of the process). Next, there is no exact analogue of Harnack's inequality in Hilbert spaces (for upper bounds see [8]); the Girsanov theorem is applicable only in very special cases where drifts take values in the Cameron–Martin spaces of the corresponding Wiener processes. Moreover, there are no full analogues of the result of Strook and Varadhan. Hence the following question arises: is the distribution of a non-degenerate diffusion process in a Hilbert space at time t positive on all open sets (at least for processes with bounded drifts)? The answer is positive for linear SDEs of the form

$$dX_t = dW_t + AX_t dt, \quad X_0 = x.$$

This equation admits an explicit solution that is a Gaussian process. However, in the general case the solution to (1) is not a Gaussian process. It needs not be even absolutely continuous with respect to a Gaussian process.

Despite the fact that this question is of considerable interest for SPDEs, only a few results in this direction are known. For some special equations (such as the stochastic Navier–Stokes equation) this question was studied by diverse methods (see [1, 17] and [18]). We also mention the paper [2], where strict positivity in the above sense was established for the invariant measure of the stochastic porous medium equation.

The problem in the general setting was considered in the book [10] for Lipschitz continuous perturbations F. The positive result for non-degenerate constant diffusion operators is obtained in [10, Theorem 7.4.2] by methods of the control theory, inspired by the ideas of Strook and Varadhan [21]. However, in this approach it is impossible to drop the assumption of the Lipschitz continuity of F.

In this paper we study the question of positivity of the distribution of non-degenerate diffusion processes on open sets with purely probabilistic methods. We consider constant non-degenerate diffusion operators and drifts that are bounded perturbations of linear operators and prove that at every positive time the distribution of such a process is positive on every ellipsoid whose axis are given by the eigenvectors of the diffusion operator. This means that the distribution has full topological support in the weaker topology in which these ellipsoids are balls. The main difference of this result from the above mentioned result in [10] is that we don't assume that the nonlinear term F is Lipschitz continuous. Instead of this, we assume that the SDE and the corresponding Kolmogorov equation have unique solutions. This is a much milder assumption since typically SDEs with non-degenerate diffusions are more regular than ODEs. Moreover, due to the fast development of the field and new results on well-posedness, this assumption is less and less restrictive. The second difference consists in using purely probabilistic methods without references to the control theory.

Let us proceed to exact statements.

Let H be a separable real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$. Fix a positive self-adjoint operator $Q\colon H \to H$ with finite trace and eigenvalues $\{q_j\}_{j\in\mathbb{N}}$. Set

$$\operatorname{tr} Q := \sum_{j=0}^{\infty} q_j < \infty.$$

We assume that

$$1 = q_1 \geq q_2 \geq \cdots > 0. \tag{4}$$

We assume that we are given an H-valued Wiener process $\big(W_t,\, t \in \mathbb{R}_+\big)$ on some probability space $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ with covariance operator Q, i.e.

$$\mathbb{E}\langle W_t, u\rangle\langle W_s, v\rangle = \min\{t, s\} \cdot \langle Qu, v\rangle.$$

Let $\big(\mathcal{F}_t,\, t \geq 0\big)$ be the filtration generated by this Wiener process. There exist an orthonormal system $\{e_j\}_{j\in\mathbb{N}}$ in H (see [9, Proposition 4.3]) and a countable set of independent one-dimensional standard Wiener

processes $(\beta_t^j,\, t\in\mathbb{R}_+),\ j\in\mathbb{N}$ on $\big(\Omega,\mathcal{F},\mathbb{P}\big)$ that are $\big(\mathcal{F}_t,\, t\geq 0\big)$-adapted such that

$$W_t=\sum_{j=1}^{\infty}\sqrt{q_j}\beta_t^j e_j, \tag{5}$$

where the series converges in L^2. Define a weighted norm on H by

$$\|x\|_Q := \langle Qx,x\rangle^{1/2}=\Big(\sum_{j=1}^{\infty}q_j x_j^2\Big)^{1/2},\quad x_j:=\langle x,e_j\rangle$$

and observe that $\|x\|_Q\leq\|x\|$ for each $x\in H$ due to (4). Given $a\in H$ and $R\in\mathbb{R}_+$, set

$$\begin{aligned}K_R(a):&=\{x\in H:\ \|x-a\|_Q\leq R\},\quad U_R(a):=\{x\in H:\ \|x-a\|\leq R\};\\ K_R(0)&=:K_R,\qquad U_R(0)=:U_R.\end{aligned}$$

The sets $K_R(a)$ will be called ellipsoids and the sets $U_R(a)$ will be called balls. The ellipsoid $K_R(a)$ contains $U_R(a)$, but is not contained in any ball $U_{R'}(a^{'})$ (contrary to the finite-dimensional case).

Let $\mathcal{B}(H)$ denote the σ-field of all Borel sets in H. Let $\mathcal{P}_\infty(H)$ denote the set of all probability measures on $(H,\mathcal{B}(H))$ with finite moments of all orders. Let $\mathcal{V}_\infty(H)$ denote the set of all H-valued random variables with finite moments of all orders. Finally, let $\mathcal{FC}_0^\infty(H)$ denote the class of all functions of the form $\phi(x)=\phi_0(x_1,\dots,x_m)$ with some $m\in\mathbb{N}$, where ϕ_0 is an infinitely smooth function with compact support in $\mathbb{R}^m$.

2 SDE with a bounded drift

First we consider the case of a bounded drift. This case is not only interesting in itself, but is also a basis for further consideration.

Suppose that an H-valued random variable η and a function $F:\ H\to H$ are given.

On the probability space $\big(\Omega,\mathcal{F},\mathbb{P}\big)$ consider the following SDE:

$$dX_t=dW_t+F(X_t)dt,\qquad X_0=\eta. \tag{6}$$

An $\mathcal{F}_t$-adapted H-valued process $\big(X_t,\, t\in\mathbb{R}_+\big)$ is said to be a strong solution to (6) if $\mathbb{P}$-a.s. for all $t\geq 0$

$$X_t=\eta+W_t+\int_0^t F(X_s)ds, \tag{7}$$

where the last integral is a Bochner integral. In the sequel we shall consider the distributions $(\mu_t)_{t\geq 0}$ of the process $\big(X_t,\, t\in\mathbb{R}_+\big)$, defined by

$$\mu_t(C)=\mathbb{P}(X_t\in C),\quad C\in\mathcal{B}(H).$$

To the diffusion process (6) we associate the Cauchy problem for its distributions

$$\partial_t\mu_t=\frac{1}{2}q_i\partial^2_{e_ie_i}\mu_t-\partial_{e_i}(b^i(x)\mu_t),\quad \mu_0=\nu=\mathrm{Law}(\eta), \tag{8}$$

where $b^i=\langle F,e_i\rangle$. Throughout the paper we assume that summation over all repeated indices is taken. A family of probability measures $(\mu_t)_{t\geq 0}$ is said to be a solution to (8) if the identity

$$\int\phi(x)d\mu_t-\int\phi(x)d\nu=\int_0^t\int\mathcal{L}\phi(x)d\mu_s ds,$$

where

$$\mathcal{L}\phi=\sum_{i=0}^{\infty}\frac{1}{2}q_i\partial^2_{e_ie_i}\phi+\sum_{i=0}^{\infty}b^i\partial_{e_i}\phi,$$

holds for all $t \geq 0$ and all test functions $\phi \in \mathcal{FC}_0^\infty(H)$.

Further we assume that

(i) η is independent of $\big(W_t,\, t \in \mathbb{R}_+\big)$ and $\eta \in \mathcal{V}_\infty\big(H\big)$;

(ii) the function F is bounded, i.e.

$$\sup_{x \in H} \|F(x)\| = F_* < +\infty.$$

(iii) The equation (6) has a strong solution $X_t, t \geq 0$ and $X_t \in \mathcal{V}_\infty(H)$ for each $t \geq 0$. The problem (8) has a unique probability solution.

Under assumption (iii) the distributions of the process X_t solve the Cauchy problem (8) (see [9, Section 14.2.2]). This one-to-one correspondence between equations enables us to switch between probability representations and measures whenever it is convenient.

Theorem 2.1. *Assume* (i), (ii) *and* (iii) *hold. Then, for any initial condition $\eta \in \mathcal{V}_\infty(H)$ and for every $T > 0$, the solution to (8) is strictly positive on every ellipsoid $K_R(a)$:*

$$\mu_T(K_R(a)) > 0, \text{ or, equivalently, } \mathbb{P}(X_T \in K_R(a)) > 0.$$

Remark 2.1. Equation (8) is meaningful for any nonnegative finite Borel initial measure ν, and then the solution is a finite nonnegative Borel measure and preserves the total mass $\nu(H)$ of the space. Hence the result of Theorem 2.1 is valid for the Cauchy problem (8) with any finite nonnegative Borel initial measure ν.

Remark 2.2. As it can be seen from the proof, in (iii) instead of existence of a strong solution it sufficies to assume only existence of a weak solution which possesses the Markov property. In regular finite-dimensional cases existence of a weak solution, together with uniqueness of distribution, ensures [15] that it is a Markov process on its probability basis. Morever, existence of weak solution is closely related to the solvability of the corresponding martingale problem, which, in it's turn, is connected to the well-posedness of the Kolmogorov equation. However, the author doesn't know any precise analogues of these results in the infinite-dimensional setting. To the author's knowledge, similar results are proved under additional assumptions like m-dissipativity of the drift or for equations with initial data from a particular class (for example, see [3]).

Proof. We split the proof into several steps.

1. We prove that for each ellipsoid $K_R(a)$, each initial distribution $\nu \in \mathcal{P}_\infty(H)$ and each $T > 0$, there exists a time $t_0 \in (0, T]$ such that at t_0 the solution to the Cauchy problem (8) is strictly positive on $K_R(a)$:

$$\mu_{t_0}(K_R(a)) = \mathbb{P}(X_{t_0} \in K_R(a)) > 0. \tag{9}$$

2. We prove that, for each ellipsoid $K_R(a)$, there exists $\tau = \tau(R) > 0$ such that for any initial distribution $\nu \in \mathcal{P}_\infty\big(H\big)$ one has

$$\mu_t(K_R(a)) = \mathbb{P}(X_t \in K_R(a)) > 0 \quad \forall t \in (0, \tau].$$

3. We prove the assertion of the theorem, i.e., that

$$\mu_t(K_R(a)) = \mathbb{P}(X_t \in K_R(a)) > 0 \quad \forall t > 0.$$

Step 1. First, let us show that for each initial measure $\nu \in \mathcal{P}_\infty\big(H\big)$ that is not Dirac's measure at zero and for each $T > 0$, there exists $t_0 \in (0, T]$ such that $\mu_{t_0}(K_R) > 0$. It suffices to prove this assertion for initial measures with $\operatorname{supp} \nu \subset U_N \backslash K_\delta$ for some $N > \delta > 0$. Indeed, assume that (9) holds for every initial measure supported in $U_N \backslash K_\delta$. The continuity of ν at zero yields that there is $\delta > 0$ such that $\nu(H \backslash K_\delta) > 0$. Since

$$U_N \subset U_{N+1} \quad \text{and} \quad \bigcup_{N=1}^{\infty} U_N = H,$$

there is an index N_0 such that $\nu(U_{N_0} \backslash K_\delta) > 0$. Define measures ν_0 and $\nu^\perp$ by

$$\nu_0(E) = \nu(E \cap (U_{N_0} \backslash K_\delta)), \quad \nu^\perp(E) = \nu(E \backslash (U_{N_0} \backslash K_\delta)).$$

Then $\nu = \nu_0 + \nu^{\perp}$. Observe that equation (8) is linear in measure, hence $\mu_t = \mu_t^0 + \mu_t^{\perp}$, where $\mu_t^0,\ \mu_t^{\perp}$ are solutions to (8) with initial measures ν_0 and $\nu^{\perp}$, respectively. By Remark 2.1, (9) holds for the family $(\mu_t^0)_{t\ge 0}$ with some $t_0 \in (0,T]$, thus

$$\mu_{t_0}(K_R) = \mu_{t_0}^0(K_R) + \mu_{t_0}^{\perp}(K_R) \ge \mu_{t_0}^0(K_R) > 0.$$

Hence we can assume from the very beginning that the initial measure ν satisfies the condition

$$\operatorname{supp}\nu \subset U_N \backslash K_\delta \quad \text{for some } N > \delta > 0.$$

In particular, ν can be an atomic measure outside zero. Fix $K_R = K_R(0)$ and $T > 0$. Let η be an H-valued random variable independent of $\big(W_t,\, t \in \mathbb{R}_+\big)$ such that $\mathrm{Law}(\eta) = \nu$.

Let us show that there exists $t_0 \in (0,T]$ such that $\mu_{t_0}(K_R) > 0$. We argue by contradiction. Suppose that this is false and $\mu_t(K_R) = 0$ for all $t \in (0,T]$. Without loss of generality we can assume that $R < \delta$ and $\mu_t(K_R) = 0$ for all $t \in [0,T]$. In particular, this means that $\mathbb{P}$-a.s. $\|X_t\| \ge \|X_t\|_Q \ge R$ for all $t \in [0,T]$.

Consider the one-dimensional stochastic process $\zeta_t = \|X_t\|^2$. It is a smooth function of the diffusion process (7) and its Itô's differential can be computed by using Itô's formula for H-valued processes (see [9, Theorem 4.32]):

$$d\zeta_t = 2\langle X_t, dW_t\rangle + (2\langle X_t, F(X_t)\rangle + \operatorname{tr} Q)dt, \quad \zeta_0 = \|\eta\|^2.$$

In order to simplify the first term in the differential, we observe that the one-dimensional stochastic process $w = (w_t,\, t \ge 0)$ given by

$$w_t = \int_0^t \frac{\langle X_s, dW_s\rangle}{\|X_s\|_Q}$$

is a continuous square-integrable $\mathcal{F}_t$-martingale and (see [9, Theorem 4.27]) its quadratic variation equals

$$\ll w_t \gg = \int_0^t \Phi_s\, ds,$$

where

$$\Phi_s = \Big(\frac{X_s}{\|X_s\|_Q} Q^{1/2}\Big)\Big(\frac{X_s}{\|X_s\|_Q} Q^{1/2}\Big)^* = \frac{1}{\|X_s\|_Q^2}\cdot \big(X_s\, Q^{1/2}\big)\big(X_s\, Q^{1/2}\big)^* = \frac{\|X_s\|_Q^2}{\|X_s\|_Q^2} = 1.$$

Hence $\ll w_t \gg = t$. Lévy's characterization of the Brownian motion (see [13, Chapter 3, Theorem 3.16]) yields that w is an $\mathcal{F}_t$-adapted Wiener process. Thus,

$$\begin{aligned} \zeta_t = \zeta_0 + \int_0^t v(\omega,s)dw_s + \int_0^t c(\omega,s)ds, \\ v(\omega,t) := 2\|X_t\|_Q, \ \ c(\omega,t) := 2\langle X_t, F(X_t)\rangle + \operatorname{tr} Q. \end{aligned} \tag{10}$$

Observe that $v(\omega,t)$ is also a progressively measurable ($\mathcal{F}_t$-adapted) process. Since

$$c(\omega,t) \le \mathrm{tr}Q + \zeta_t + \|F\|_\infty^2 =: \lambda + \zeta_t,$$

by the assumption $\zeta_0 \le N^2$ we have

$$\zeta_t \le \zeta_0 + \int_0^t v(\omega,s)dw_s + \int_0^t (\lambda + \zeta_s)ds \le (N^2 + T\lambda) + \int_0^t v(\omega,s)dw_s + \int_0^t \zeta_s ds. \tag{11}$$

Letting

$$\Psi_t := (N^2 + T\lambda) + \int_0^t v(\omega,s)dw_s,$$

we obtain

$$\zeta_t \le \Psi_t + \int_0^t \zeta_s ds. \tag{12}$$

Multiplying by e^{-t}, we obtain

$$\frac{d}{dt}\Big(e^{-t}\cdot\int_0^t \zeta_s ds\Big)\le e^{-t}\Psi_t,\quad \text{hence}\quad \int_0^t \zeta_s ds\le\int_0^t e^{t-s}\Psi_s ds.$$

Plugging this estimate into (11), we arrive at

$$0\le\zeta_t\le\Psi_t+\int_0^t e^{t-s}\Psi_s ds\le C(N,T)+\int_0^t v(\omega,s)dw_s+\int_0^t e^{t-s}\int_0^s v(\omega,r)dw_r ds, \tag{13}$$

where $C(N,T):=(N^2+T\lambda)(1+Te^T)>0$. Next, by the integration by parts formula (see [19, Ex. 4.3]), we have

$$\int_0^t e^{t-s}\int_0^s v(\omega,r)dw_r ds=e^t\Big(-e^{-t}\int_0^t v(\omega,r)dw_r\Big)+e^t\int_0^t e^{-s}d\int_0^s v(\omega,r)dw_r$$
$$=-\int_0^t v(\omega,r)dw_r+e^t\int_0^t e^{-s}v(\omega,s)dw_s,$$

hence (13) implies that for $t\in[0,T]$

$$\int_0^t e^{-s}v(\omega,s)dw_s\ge -C(N,T)\cdot e^{-t}\ge -C(N,T). \tag{14}$$

By our assumption $v\ge 2R$. Fix an arbitrary $t^*\in(0,t)$. Define a random change of time

$$z_t:=\int_0^t e^{-2s}v^2(\omega,s)ds\ge t\cdot(2R)^2e^{-2T}. \tag{15}$$

For each $\gamma\ge 0$, set $\tau_\gamma:=\inf\{s\ge 0:\ z_s=\gamma\}$. The paths of the process z_t are continuous and the process is bounded from below according to (15), hence τ_γ is a stopping time with respect to the filtration $(\mathcal{F}_t,\, t\ge 0)$. Moreover, $\mathbb{P}(\tau_\gamma<+\infty)=1$ and $\tau_\gamma<t^*$ for each $\gamma<t^*\cdot(2R)^2e^{-2T}$ with $\mathbb{P}$-probability 1. The change of time theorem ([12, Chapter 1, Par. 4, Theorem 3]) implies that the stochastic process $y=(y_\gamma,\,\gamma\ge 0)$ given by

$$y_\gamma:=\int_0^{\tau_\gamma} e^{-s}v(\omega,s)dw_s$$

is a Wiener process with respect to the filtration $(\mathcal{F}_{\tau_\gamma},\,\gamma\ge 0)$. In particular, the random variable y_γ has a strictly positive distribution density on the real line. On the other hand, $\int_0^t e^{-s}v(\omega,s)dw_s$ is an $\mathcal{F}_t$-martingale. It is well-known (see, for example, [22, Paragraph 7.2, Microtheorem 3]) that the martingale property holds not only for deterministic times, but also for bounded stopping times: $\mathbb{P}$-a.s. one has

$$y_\gamma=\int_0^{\tau_\gamma} e^{-s}v_s dw_s=\int_0^{\tau_\gamma\wedge t} e^{-s}v_s dw_s=\mathbb{E}\Big(\int_0^t e^{-s}v_s dw_s|\,\mathcal{F}_{\tau_\gamma}\Big)\ge -C(N,T),$$

since $\tau_\gamma<t^*<t$. This contradiction means that there exists $t_0\in(0,T]$ such that $\mu_{t_0}(K_R)>0$.

Let us now proceed to non-centered ellipsoids. Fix $K_R(a)$ with a center $a\in H$. Let us show that there is $t_0\in(0,T]$ such that the solution to (8) is positive on $K_R(a)$ for very initial measure $\nu\ne\delta_a$. Fix $\nu\ne\delta_a$.

Consider the shift $L^a\colon\ H\to H$ defined by

$$L^a x=x+a.$$

We recall that the image of a measure ρ under the mapping L^a is the measure $L^a_*\rho$ defined by $L^a_*\rho(E)=\rho(L^a(E)))$ for each measurable set $E\subset H$. Then it follows from the definition that $L^a(K_R)=K_R(a)$ and the measures $\sigma_t=L^a_*\mu_t$ satisfy the equation

$$\partial_t\sigma_t=\frac{1}{2}q_i\partial^2_{e_ie_i}\sigma_t-\partial_{e_i}(b^i(x-a)\sigma_t),\quad \sigma_0=L^a_*\nu\ne\delta_0,$$

where $b^i(\cdot - a) = \langle F(\cdot - a), e_i\rangle$. The drift term $F(\cdot - a)$. Therefore, by the assertion for centered balls proved above in the case $\sigma_0 \neq \delta_0$, there exists $t_0 \in (0, T]$ such that

$$\mu_{t_0}(K_R(a)) = \mu_{t_0}(L^a(K_R)) \stackrel{def}{=} L^a_* \mu_{t_0}(K_R) > 0.$$

To complete the proof of this step, we consider $K_R(a)$ and $\nu = \delta_a$. Note that for $\varepsilon > 0$ small enough

$$K_{R/2}(a + \bar{\varepsilon}) \subset K_R(a), \quad \bar{\varepsilon} = \varepsilon \cdot e_1 \in H.$$

Indeed, if $(x_1 - a_1 - \varepsilon)^2 + \sum_{j=2}^{\infty} q_j (x_j - a_j)^2 \leq R^2/4$, then

$$\sum_{j=1}^{\infty} q_j (x_j - a_j)^2 \leq 2(x_1 - a_1 - \varepsilon)^2 + 2\varepsilon^2 + \sum_{j=2}^{\infty} q_j (x_j - a_j)^2 \leq R^2/2 + 2\varepsilon^2 \leq R^2$$

for $\varepsilon^2 \leq R^2/4$. But (9) has already been proved for $K_{R/2}(a + \bar{\varepsilon})$ and $\nu = \delta_a$, i.e. $\mu_{t_0}(K_{R/2}(a + \varepsilon)) > 0$ for some $t_0 \in (0, T]$. By additivity $\mu_{t_0}(K_R(a)) \geq \mu_{t_0}(K_{R/2}(a + \bar{\varepsilon})) > 0$.

Step 2. Let us prove that for every ellipsoid $K_R(a)$, there exists $\tau = \tau(R) > 0$, depending only on R and sup-norm of F, such that *for any* initial distribution $\nu \in \mathcal{P}_\infty(H)$ one has

$$\mathbb{P}(X_t \in K_R) > 0 \quad \text{for all } t \in (0, \tau(R)],$$

where X_t solves (6).

The idea of the proof is quite simple: if the process with any initial distribution at some time t_0 hits a small ellipsoid with positive probability, then with positive probability it stays in a larger ellipsoid during some time, and this time is determined by the parameters of the ellipsoid. But it has already been proven that during every small interval of time the process X_t hits every fixed ellipsoid (with positive probability) at least once. The combination of these facts yields the assertion of Step 2. Let us proceed to rigorous proofs.

Fix $X_0 \in \mathcal{V}_\infty(H)$ and $K_R(a)$. Set

$$\tau(R) := R \cdot \Big(6 \cdot (1 + \sup_{x \in H} \|F(x)\|_Q)\Big)^{-1}.$$

Lemma 2.1. *Assume that* $supp\,\nu \subset K_{R/2}(a)$. *Then*

$$\mathbb{P}(X_t \in K_R(a)) > 0 \quad \textit{for all } t \in (0, \tau(R)]. \tag{16}$$

Proof of Lemma 2.1. Recall that

$$X_t = X_0 + W_t + \int_0^t F(X_s) ds.$$

Obviously, it suffices to show that for all $t \in (0, \tau(R)]$

$$\mathbb{P}\Big(\|X_t - X_0\|_Q > R/2\Big) < 1.$$

This follows from the properties of H-valued Wiener processes and the definition of $\tau(R)$. Indeed,

$$\begin{aligned}\mathbb{P}\Big(\Big\|X_t - X_0\Big\|_Q > \frac{R}{2}\Big) &= \mathbb{P}\Big(\Big\|W_t + \int_0^t F(X_s) ds\Big\|_Q > \frac{R}{2}\Big) \\ &\leq \mathbb{P}\Big(\|W_t\|_Q > \frac{R}{4}\Big) + \mathbb{P}\Big(\Big\|\int_0^t F(X_s) ds\Big\|_Q > \frac{R}{4}\Big).\end{aligned} \tag{17}$$

By the properties of the Bochner integral and the definition of $\tau(R)$ we have

$$\Big\|\int_0^t F(X_s) ds\Big\|_Q \leq \int_0^t \|F(X_s)\|_Q ds \leq \tau(R) \cdot \sup_{x \in H} \|F(x)\|_Q \leq \frac{R}{6} < \frac{R}{4},$$

i.e. the second probability on the right-hand side of (17) equals zero. Hence

$$\mathbb{P}(\|X_t - X_0\|_Q > \frac{R}{2}) \le \mathbb{P}(\|W_t\|_Q > \frac{R}{4}) \le \mathbb{P}(\|W_t\| > \frac{R}{4}).$$

The distribution of W_t at time t is a centered H-valued Gaussian random variable with variance $t \cdot Q$. By [7, Theorem 3.5.1] the probability on the right-hand side of the last inequality is strictly less than 1. This completes the proof of Lemma 2.1. $\square$

Let us return to the proof of Step 2. Fix $\delta \in (0, \tau(R))$. According to Step 1 there exists a time $t_0 \in (0, \delta)$ such that $\mu_{t_0}(K_{R/2}(a)) > 0$. By the Markov property

$$X_t = X_{t_0} + W_t^1 + \int_{t_0}^t F(X_s)ds, \quad t \ge t_0$$

where $W^1 = (W_t - W_{t_0},\, t \ge t_0)$ is also a Q-Wiener process. By our choice of t_0 we have

$$\mathbb{P}(X_{t_0} \in K_{R/2}(a)) = \mu_{t_0}(K_{R/2}(a)) > 0.$$

Arguing similarly to Step 1 and applying Lemma 2.1, we obtain

$$\mathbb{P}(X_t \in K_R(a)) = \mu_t(K_R(a)) > 0 \quad \forall t \in [t_0, t_0 + \tau(R)].$$

In particular, this holds for all $t \in [\delta, \tau(R)]$, but δ is an arbitrary number in $(0, \tau(R)]$, hence

$$\mathbb{P}(X_t \in K_R(a)) = \mu_t(K_R(a)) > 0 \quad \text{for all } t \in (0, \tau(R)]. \tag{18}$$

Step 3. Fix an arbitrary time M. Split the interval $[0, M]$ into $n := [M/\tau(R)]$ parts, where $\tau(R)$ is defined by (16):

$$[0, M] = \bigcup_{i=0}^{n-1} [s_i, s_{i+1}], \quad s_j = j \cdot \tau(R),\ j = 0, \ldots, n-1, \quad s_n = M.$$

By the previous step, for any initial data $\eta \in \mathcal{V}_\infty(H)$, the assertion of Theorem 2.1 holds on $[0, \tau(R)] \equiv [s_0, s_1]$, i.e. (18). Similarly to the Step 2, we have

$$X_t = X_{s_1} + W_t^2 + \int_{s_1}^t F(X_s)ds, \quad t \ge s_1.$$

Application of the result of Step 2 gives that $\mathbb{P}(X_t \in K_R(a)) = \mu_t(K_R(a)) > 0$ for $t \in (s_1, s_2]$. By induction we get

$$\mu_t(K_R(a)) > 0 \quad \text{for all } t \in (0, M].$$

This completes the proof of Theorem 2.1. $\square$

Remark 2.3. If $u(\omega, t) = e^{-t} v(\omega, t)$ is not separated from zero, then, generally speaking, (14) does not yield a contradiction. This can be shown by a simple example (suggested by A.A. Novikov). Consider $u(\omega, t) = \exp\{w_t - t/2\} > 0$ $\mathbb{P}$-a.s., where w_t is a standard Wiener process on the real line. Obviously, there is no positive R such that $\mathbb{P}$-a.s. $u(\omega, s) \ge R$. Itô's formula implies

$$u(\omega, t) = 1 + \int_0^t u(\omega, s) dw_s > 0, \quad \text{hence} \int_0^t u(\omega, s) dw_s > -1 \quad \mathbb{P}-\text{a.s.}$$

Remark 2.4. The assumption (iii) is fulfilled, for example, if F is Lipschitz continuous. Equation (6) has a unique strong solution $X_t, t \ge 0$ due to [9, Theorem 7.2] and $X_t \in \mathcal{V}_\infty(H)$ for each $t \ge 0$. The problem (8) has a unique probability solution by virtue of [5, Theorem 1] and [4, Theorem 2.1]. However, Theorem 2.1 is in a sense stronger than [10, Theorem 7.4.2], mentioned in the Introduction, where irreducibility of the corresponding semigroup is demonstrated, because it does not require any continuity of the nonlinear perturbation.

3 SDE with unbounded drift

We now proceed to the general case – SDE (1) with an unbounded self-adjoint negative linear operator A:

$$dX_t = dW_t + (AX_t + F(X_t))dt, \quad X_0 = \eta. \tag{19}$$

Here, as above, $\big(W_t,\, t \in \mathbb{R}_+\big)$ is a Q-Wiener process on $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ with the natural filtration $(\mathcal{F}_t,\, t \geq 0)$. Set $B(x) = Ax + F(x)$.

Let us now recall the concept of variational solution (see [20]).

Consider the Banach space $V := D((-A)^{1/2})$ equipped with the graph norm of $(-A)^{1/2}$ and its dual space V^*. Then (V, H, V^*) is a Gelfand triple, i.e. $V \subset H \subset V^*$ and the embeddings are continuous and dense. Let us consider the Friedrichs extension A_1 of A. Then $A_1\colon V \to V^*$ and A_1 is also a densely defined negative self-adjoint operator (see, for example, [14, Theorem 2.23]). Set $B_1(\cdot) := A_1 + F(\cdot)\colon V \to V^*$. For notational simplicity, further we omit indices, and A will denote not only the operator, but also its Friedrichs extension, and also $B(\cdot) = A + F(\cdot)$.

A continuous H-valued $\mathcal{F}_t$-adapted process $X = \big(X_t,\, t \in [0,T]\big)$ is called a variational solution to (19) if for its $dt\times\mathbb{P}$-equivalence class $\hat{X}$ with some $\alpha \geq 1$ we have $\hat{X} \in L^\alpha([0,T]\times\Omega, dt\times\mathbb{P}; V)\cap L^2([0,T]\times\Omega, dt\times\mathbb{P}; H)$ and $\mathbb{P}$-a.s.

$$X_t = \eta + W_t + \int_0^t B(\bar{X}_s)ds, \quad t \in [0,T], \tag{20}$$

where $\bar{X}$ is any $\mathcal{F}_t$-adapted V-valued $dt \times \mathbb{P}$-version of $\hat{X}$. Moreover, the integrand in (20) is automatically H-valued (see, for example, [20, Remark 4.2.2]). Below we set $\alpha = 2$.

Along with assumptions (i) and (ii) from the previous section, we shall need the following assumptions:

(iii') The problem (8) has a unique probability solution. The equation (19) has a variational solution (see [16]) and

$$\mathbb{E} \sup_{t\in[0,T]} \|X_t\|^2 < +\infty. \tag{21}$$

(iv) The domain $D(A) \subset H$ of the linear operator A is dense in H and A is self-adjoint and negative (i.e. $\langle Ax, x\rangle \leq -\varepsilon\|x^2\|$ for some $\varepsilon > 0$ and all $x \in H$).

The Hille–Yosida theorem (see, for example, [11, Theorem 2.6]) states that any linear operator A with properties (iv) generates a contracting strongly continuous semigroup S_t, $t \in \mathbb{R}_+$ of linear transformations of H.

A continuous $\mathcal{F}_t$-adapted H-valued process $X = \big(X_t,\, t \in [0,T]\big)$ is said to be a mild solution to (19) (see, for example, [9, 20]) if $\mathbb{P}$-a.s. for all $t \in [0,T]$ one has

$$X_t = S_t\eta + \int_0^t S_{t-s}I\, dW_s + \int_0^t S_{t-s}F(X_s)ds. \tag{22}$$

Here I is the identity operator on H; the last integration is in Bochner's sense.

The distributions of the process X_t solve (8) with $b^i = \langle B, e_i\rangle$ (see [9, Section 14.2.2]). As above, this one-to-one correspondence enables us to consider measures in placed of processes and vice versa, whenever this is convenient.

The main result of this section is the following theorem.

Theorem 3.1. *Assume that* (i), (ii), (iii$'$) *and* (iv) *hold. Then, for any initial condition $\eta \in \mathcal{V}_\infty(H)$ and for every $t \in (0,T]$, the solution to (8) is strictly positive on each ellipsoid $K_R(a)$:*

$$\mu_t(K_R(a)) > 0, \ \textit{or, equivalently, } \mathbb{P}(X_t \in K_R(a)) > 0.$$

Proof. The proof mainly repeats the proof of Theorem 2.1. We consider only the steps affected by the addition of the linear term.

Arguing similarly to Step 1 of the proof of Theorem 2.1 and applying Itô's formula for variational solutions (see [20, Theorem 4.2.5]), we obtain the following expression for the process $\zeta_t = \|X_t\|^2$:

$$\zeta_t = \zeta_0 + \int_0^t 2\|X_s\|_Q dw_s + \int_0^t (2\langle X_s, F(X_s)\rangle + \operatorname{tr} Q + 2\langle AX_t, X_t\rangle)ds$$
$$\leq \zeta_0 + \int_0^t 2\|X_s\|_Q dw_s + \int_0^t (2\langle X_s, F(X_s)\rangle + \operatorname{tr} Q)ds.$$

where we used the estimate $\langle Ax, x\rangle \leq 0$. Similarly to the derivation of the bound (14), we obtain

$$\int_0^t e^{-s} v(\omega, s) dw_s \geq -C, \qquad v(\omega, s) := 2\|X_s\|_Q.$$

Step 1 is completed in exactly the same way as in proof of Theorem 2.1. Next, we observe that the structure of the drift term in the proof of Theorem 2.1 has only been used in Lemma 2.1. Therefore, to complete the proof of Theorem 3.1 it suffices to prove an analogue of Lemma 2.1 in the case $A \neq 0$. Fix $X_0 \in \mathcal{V}_\infty\big(H\big)$ and $K_R(a)$. Let $\nu = \mathrm{Law}(X_0)$. Set

$$\tau(R) := R \cdot (6 \cdot (1 + \sup_{x\in H} \|F(x)\|))^{-1}.$$

Lemma 3.1. *Suppose that X_0 is independent of $\big(W_t,\, t \in \mathbb{R}_+\big)$ and supp $\nu \subset K_{R/2}(a)$. Then*

$$\mathbb{P}\Big(X_t \in K_R(a)\Big) > 0 \quad \textit{for all } t \in (0, \tau(R)]. \tag{23}$$

Proof of Lemma 3.1. Note that the variational solution X_t is also a mild solution to (1) (see [20, F.0.5, F.0.6]), i.e.

$$X_t = S_t X_0 + \int_0^t S_{t-s} I\, dW_s + \int_0^t S_{t-s} F(X_s) ds.$$

Clearly, it suffices to prove that for all $t \in (0, \tau(R)]$

$$\mathbb{P}(\|X_t - X_0\|_Q > R/2) < 1, \text{ if } \quad \mathrm{Law}(X_0) = \nu.$$

We have

$$\mathbb{P}\Big(\|X_t - X_0\|_Q > \frac{R}{2}\Big) \leq \mathbb{P}\Big(\Big\|(S_t - I)X_0 + \int_0^t S_{t-s} I\, dW_s\Big\|_Q > \frac{R}{4}\Big)$$
$$+ \mathbb{P}\Big(\Big\|\int_0^t S_{t-s} F(X_s) ds\Big\|_Q > \frac{R}{4}\Big). \tag{24}$$

Since the semigroup S_t is contracting,

$$\Big\|\int_0^t S_{t-s} F(X_s) ds\Big\|_Q \leq \Big\|\int_0^t S_{t-s} F(X_s) ds\Big\| \leq \int_0^t \|S_{t-s} F(X_s)\| ds \leq \tau(R) \cdot \sup_{x\in H} \|F(x)\| < \frac{R}{4},$$

i.e. the second probability on the right-hand side of (24) is zero. Thus,

$$\mathbb{P}\Big(\|X_t - X_0\|_Q > \frac{R}{2}\Big) \leq \mathbb{P}\Big(\Big\|(S_t - I)X_0 + \int_0^t S_{t-s} I\, dW_s\Big\|_Q > \frac{R}{4}\Big)$$
$$\leq \mathbb{P}\Big(\|(S_t - I)X_0 + \int_0^t S_{t-s} I\, dW_s\| > \frac{R}{4}\Big). \tag{25}$$

The process $W_A = (W_A(t),\, t \geq 0)$ given by $W_A(t) := \int_0^t S_{t-s} I\, dW_s$ is called a stochastic convolution. Since

$$\int_0^T \operatorname{tr} S(r) Q S^*(r) dr = \operatorname{tr} \int_0^T \|S(r)\|_Q^2 dr < \infty,$$

W_A is an $\mathcal{F}_t$-adapted Gaussian random variable, continuous in mean square, with the non-degenerate covariance operator $\int_0^t \|S(r)\|_Q^2 dr$ (see [9, Theorem 5.2]). It can be easily seen that $(S_t - I)X_0$ and $W_A(t)$ are independent random variables. By the convolution formula

$$\mathbb{P}(\|(S_t - I)X_0 + \int_0^t S_{t-s} I\, dW_s\| \leq \frac{R}{4}) = \int_H \rho_t(U_{R/4}(0) - w)\sigma_t(dw), \tag{26}$$

where $\sigma_t = \mathrm{Law}(S_t - I)X_0$ and $\rho_t = \mathrm{Law}(W_A(t))$. But the integrand is strictly positive by the properties of the Gaussian random variable ρ_t (see [7, Theorem 3.5.1]), and σ_t is a probability measure, hence (26) is a strictly positive quantity. Therefore, the right-hand side of (25) is strictly less than 1. This completes the proof of Lemma 3.1 and Theorem 3.1. □ □

Acknowledgements

The author is grateful to V.I. Bogachev, A.A. Novikov, A.Yu. Veretennikov and A.D. Manita for fruitful discussions. The work was partially supported by RFBR grants 14-01-00237 and 14-01-00319.

References

[1] A. Agrachev, S. Kuksin, A. Sarychev, and A. Shirikyan. *On finite-dimensional projections of distributions for solutions of randomly forced 2d navier–stokes equations*, Annales de l'Institut Henri Poincare (B) Probability and Statistics, **43** (2007), 4, p. 399 – 415.

[2] V. Barbu, V.I. Bogachev, G. Da Prato, and M. Röckner. *Weak solutions to the stochastic porous media equation via kolmogorov equations: the degenerate case*, Journal of Functional Analysis, **237** (2006), 1, p. 54–75.

[3] L. Beznea, N. Boboc and M. Röckner. *Markov processes associated with L^p-resolvents and applications to stochastic differential equations on Hilbert space*, Journal of Evolutional Equations, **6** (2006), 4, p. 745-772.

[4] V.I. Bogachev, G. Da Prato, M. Röckner, and S.V. Shaposhnikov. *An analytic approach to infinite-dimensional continuity and Fokker–Planck–Kolmogorov equations*, Annali della Scuola Normale Superiore di Pisa, Classe di Scienze, **5**, vol. XIV (2015), p. 983–1023.

[5] V.I. Bogachev, G. Da Prato, and M. Röckner. *Parabolic equations for measures on infinite-dimensional spaces*, Doklady Mathematics, **78** (2008), 1.

[6] V.I. Bogachev, N.V. Krylov, and M. Röckner. *On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions*, Communications in Partial Differential Equations, **26** (2001), 11, p. 2037–2080.

[7] V.I. Bogachev. *Gaussian measures*, AMS, 1998.

[8] G. Da Prato, M. Röckner, and F.-Y. Wang. *Singular Stochastic Equations on Hilbert Spaces: Harnack Inequalities for their Transition Semigroups*, J. Funct. Anal., **257**(2009), 4.

[9] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, Cambridge University Press Cambridge, 1992.

[10] G. Da Prato and J. Zabczyk. *Ergodicity for Infinite Dimensional Systems*, Cambridge University Press Cambridge, 1996.

[11] S.N. Ethier and T.G. Kurtz. *Markov processes : characterization and convergence*, Wiley series in probability and mathematical statistics. J. Wiley and Sons, New York, Chichester, 1986.

[12] I.I. Gihman and A.V. Skorohod. *Stochastic Differential Equations*, Springer-Verlag, New York, 1972.

[13] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of Graduate Texts in Mathematics, Springer-Verlag New York, 1998.

[14] T. Kato. *Perturbation Theory for Linear Operators*, Classics in Mathematics. Springer-Verlag, Berlin-Heidelberg, 1995.

[15] N.V. Krylov. *On the selection of a Markov process from a system of processes and the construction of quasi-diffusion processes*, Mathematics of the USSR-Izvestiya, **7:3**(1973), p. 691709.

[16] N.V. Krylov and B.L. Rozovskii. *Stochastic evolution equations*, Current problems in Mathematics, VINITI, page 71, 1979.

[17] S. Kuksin and A. Shirikyan. *Mathematics of Two-Dimensional Turbulence*, Cambridge University Press, 2012.

[18] J.C. Mattingly and E. Pardoux. *Malliavin Calculus for the Stochastic 2D Navier–Stokes Equation*, arXiv:math/0407215.

[19] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*, Springer, 2014.

[20] C. Prévôt and M. Röckner. *A Concise Course on Stochastic Partial Differential Equations*, volume 1905 of Lecture Notes in Mathematics, Springer, Berlin-Heidelberg, 2007.

[21] D.W. Stroock and S.R.S. Varadhan. *On the support of diffusion processes with applications to the strong maximum principle*, In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 3: Probability Theory*, pages 333–359, Berkeley, Calif., 1972. University of California Press.

[22] A.D. Wentzell. *A Course in the Theory of Stochastic Processes*, McGraw-Hill, New York, 1981.